PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 33 August 15, 1947 Number 8

*A SUPPRESSOR IN NEUROSPORA AND ITS USE AS EVIDENCE FOR ALLELISM**

BY MARY B. HOULAHAN AND HERSCHEL K. MITCHELL

WILLIAM G. KERCKHOFF LABORATORIES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

Communicated June 10, 1947

In Drosophila recessive mutations have been observed which suppress the effect of specific mutations at other loci. Thus an individual, homozygous for a suppressor gene in the mutant form and for the mutant gene upon which the suppressor acts, is phenotypically wild type, or nearly so. Suppressors of black,[1] purple,[2, 3] sable[4] and vermilion[4, 5] are among those which have been reported. In the vermilion case there is evidence[5] which shows that the action of the suppressor is to restore v^+ substance (kynurenine), the formation of which is prevented when the mutant gene vermilion is present in the homozygous condition.

A similar situation which has been found in Neurospora will be reported here. The effect of a mutation which prevents the synthesis of pyrimidine is partly suppressed by the presence of a mutant gene at a different locus. It seems logical to assume that the presence of the second mutation results in the formation of an intermediate in pyrimidine synthesis which is lacking in the pyrimidineless strain. Evidence that the new mutation is not due simply to duplication of the pyrimidineless locus will be presented.

Results of studies of combinations of the suppressor with five pyrimidineless strains have supplemented other evidence regarding relationships among the mutant genes concerned. Genetic difference is readily demonstrated in Neurospora[6] but evidence for allelism is less convincing, because neither the absence of crossing over, nor the heterocaryon test[7] is conclusive. For this reason additional evidence, obtained by use of the suppressor, that three pyrimidineless strains carry allelic genes, is welcome, particularly since these strains show marked physiological differences, and therefore constitute the first series of multiple alleles to be described in Neurospora. This series may prove useful, since one would expect biochemical studies to be simpler in Neurospora than in more complex

organisms such as corn, Drosophila and mammals, in which multiple alleles have been studied in detail.

Forty-five independent occurrences of mutation which results in pyrimidine deficiency have been recorded in Neurospora. These were obtained from x-ray and ultra-violet treatment as described by Beadle and Tatum.[8] The deficient strains utilize uracil or the nucleosides and nucleotides of uracil and cytosine, as will be reported in detail elsewhere.[9] One of the five strains to be discussed here, number 263, has been described by Loring and Pierce.[10]

Crosses to Determine Genetic Relationships.—Isolation numbers of the strains considered are 263, 38502, 37301, 37815 and 67602, but for convenience they will be referred to here as 1, 2, 3a, 3b and 3c. Each strain behaves in backcrosses as if the difference from the parent wild type were due to modification of one gene. That is, when any one is back-crossed to the parent strain, and the eight ascospores from each of a number of asci are isolated and tested using techniques reviewed by Beadle,[6] four of the eight spores always prove to be like the mutant parent and four like the wild-type parent.

TABLE 1

CROSSES BETWEEN MUTANT STRAINS

STRAINS CROSSED	NUMBER OF ASCI OBSERVED	ASCI WITH 1 PAIR OF WILD-TYPE SPORES	ASCI WITH 2 PAIRS OF WILD-TYPE SPORES	ASCI WITH 4 PAIRS OF MUTANT SPORES
1 × 3a	37	5	0	32
1 × 3b	94	8	0	86
1 × 2	17	11	0	6
1 × 3c	53	5	0	48
3a × 3b	59	0	0	59
3a × 2	17	10	0	7
3a × 3c	15	0	0	15
3b × 2	19	6	1	12
3b × 3c	100	0	0	100
2 × 3c	21	7	0	14

Table 1 gives results of intercrosses among the five strains. (In all cases here only asci are included from which at least one member of each spore pair germinated.) From these results it appears that 3a, 3b and 3c are alike genetically, while 1 and 2 are different from these three and from each other. This is evident since from crosses among the first three, no asci were observed containing wild-type spores, while all crosses involving 1 and 2 gave some asci of this type. In order to confirm evidence of genetic difference it was demonstrated that an ascus containing wild-type spores also contained spores carrying two mutant genes. This was done by outcrossing a mutant strain derived from such a spore to wild type and recovering asci containing six or eight mutant spores. If an ascus from the

intercross contains four wild-type spores, then obviously the other four spores must carry both mutant genes, but if the ascus contains two wild-type spores then only two of the six mutant spores will give rise to double mutants. If only asci of the latter type are obtained from an intercross it would be necessary to cross to wild type a strain derived from one member of each of the three pairs of mutant spores, were it not often possible to distinguish the double mutants, partly on the basis of the position of spores in the ascus and partly because of slight differences in growth habit. Results from crosses of double mutants to wild type appear in table 2.

It is interesting that the three loci represented are on the same chromosome, as can be seen from the data in tables 1 and 2. The distances from the centromere are, roughly, 12, 39 and 17 units for 1, 2 and 3, respectively.

TABLE 2

CROSSES BETWEEN DOUBLE MUTANTS AND WILD TYPE

STRAINS CROSSED	NUMBER OF ASCI OBSERVED	ASCI WITH 1 PAIR OF WILD-TYPE SPORES	ASCI WITH 2 PAIR OF WILD-TYPE SPORES	ASCI WITH 4 PAIRS OF MUTANT SPORES
1, 3a × wild	34	8	26	0
1, 3b × wild	19	3	16	0
1, 2 × wild	12	7	5	0
1, 3c × wild	27	2	25	0
3a, 2 × wild	1	1	0	0
3b, 2 × wild	10	8	1	1
2, 3c × wild	6	4	2	0

A Suppressor of 3a.—A strain of 3a, which had begun to grow in the absence of pyrimidine as if a back mutation had occurred, was tested by crossing it to wild type. Two types of asci would be expected if back mutation had taken place, one containing eight wild-type spores, arising from nuclei carrying the back mutation, and the other, from nuclei carrying the mutant gene unchanged, containing four wild-type and four mutant spores. These two types were found, but a third type appeared as well, containing six wild-type and two mutant spores. This third type of ascus cannot be explained on the basis of back mutation at the 3a locus, but, rather, it is necessary to assume that, at a different locus, a mutation had occurred which, in some way, suppresses the effect of the mutation at the 3a locus. If this assumption were correct then all three types of asci would be expected, if there were no close linkage. The second type could result either from nuclei which did not carry the new mutation or from failure of the two mutant genes to combine in any spore pair. The first type could result from combination of the two mutations in four spores, and the third from combination in two spores.

From an ascus with eight wild-type appearing spores one member of each spore pair was crossed to wild type. Two of these crosses gave only

wild-type progeny, while from the other two, the three types of asci just described were recovered, demonstrating that the assumption made above is correct. As a further check, an ascus containing four wild-type and four mutant spores was selected from one of the above two crosses. A wild-type spore from such an ascus should carry the suppressor gene. This proved to be the case, for when a strain from one of these spores was crossed to an unchanged strain of 3a the three ascus types described above were recovered.

TABLE 3

CROSSES BETWEEN MUTANTS AND THE SUPPRESSOR[a]

STRAINS CROSSED	NUMBER OF ASCI OBSERVED	ASCI WITH 4 PAIRS OF WILD-TYPE SPORES	ASCI WITH 3 PAIRS OF WILD-TYPE SPORES	ASCI WITH 2 PAIRS OF WILD-TYPE SPORES
1 × S	19	0	0	19
3a × S	6	1	1	4
3a-S × wild	38	10	22	6
3b × S	19	2	10	7
2 × S	17	0	0	17
3c × S	11	3	2	6
3c-S × wild	14	4	7	3

[a] In some cases germination of spores from crosses between mutants and the suppressor was very poor. For this reason data from crosses of mutant-suppressor combinations to wild-type are included.

The Effect of the Suppressor on Other Pyrimidineless Mutants.—Results, given in table 3, from crosses between the suppressor strain and the remaining four mutants show that the suppressor has no effect when in combination with 1 and 2, since each ascus recovered contained four mutant and four wild-type spores. In order to show that these two mutants had not simply failed to combine with the suppressor each was crossed to a strain carrying both 3a and the suppressor gene in the mutant form. The appearance, among the progeny of both crosses, of asci containing six or eight mutant spores demonstrates that in these asci the two mutants had indeed combined with the suppressor and had not been affected by it. This follows from the fact that in a cross involving the suppressor and two mutants, both of which were capable of being suppressed, no asci could occur containing more than four mutant spores.

The data in table 3 show that 3b and 3c, as well as 3a, are suppressed. This fact is taken as evidence that mutation at the same locus has occurred in the three strains.

Growth Requirements of the Mutants.—The three mutants which behave as alleles all have much the same requirement for uridine or hydrolyzed nucleic acid if they are allowed to grow at 35°C., but at 25°C. there are striking differences (Table 4). At this temperature 3a requires nearly the same amount of uridine or hydrolyzed nucleic acid as it does at 35°C.,

while 3c needs only about one sixth as much and 3b grows normally in the absence of any source of pyrimidine. Hence it would appear that the genetic block is complete in the three strains at 35°C., and in 3a at 25°C. also, but is only partial in 3c at 25°C. and seemingly non-existent in 3b at this temperature.

TABLE 4

GROWTH REQUIREMENT OF MUTANT STRAINS

STRAIN	HYDROLYZED NUCLEIC ACID* REQUIRED FOR 1/2 MAXIMUM GROWTH (MG.) 25°C.	35°C.
3a	3.3	3.15
3b	0	2.4
3c	0.38	2.3

* Ribonucleic acid in 1 N NaOH, 24 hours at 25°C.

Strain 2 has not been thoroughly tested, but the data available indicate that its requirements are similar to those of 3c. Strain 1 has much the same requirement for uridine as 3b, but it differs from the other four strains in that it can utilize orotic acid as a substitute for uridine. Further studies along these lines will be reported elsewhere.[9]

Growth Characteristics of the Suppressed Mutant.—Preliminary experiments have shown that growth of a strain carrying the mutant gene 3a and the suppressor is not completely normal in the absence of pyrimidine. The dry weight of mycelium obtained after four days growth at 25°C. in 20 ml. of medium is about 40 mg. while that obtained using a wild-type strain under the same conditions is about 80 mg. Normal growth of the suppressed mutant takes place if any one of the following supplements is added to the medium: 0.5 mg. of uridine, 3 mg. of hydrolyzed nucleic acid, 10 mg. of lysine or 10 mg. of histidine. If the medium contains 0.5 microgram of arginine no growth takes place, although arginine in concentrations as high as 10 mg. in 20 ml. does not affect the growth of the wild type. Citrulline and ornithine are also inhibitory but are far less so than arginine. The quantities of these three compounds which, in 20 ml. of medium will allow half-maximum growth are as follows: arginine 0.15 microgram, citrulline 230 micrograms and ornithine 400 micrograms. Inhibition by arginine can be overcome by hydrolyzed nucleic acid or lysine, but not by histidine. The five mutant strains without the suppressor, when grown on a medium containing a quantity of hydrolyzed nucleic acid which allows half-maximum growth, or less, are not affected by the addition of 5 micrograms of arginine, but show a slight inhibition if 10 mg. of histidine are added.

No interpretation which is consistent with all of these facts has been made as yet, but further experiments are in progress. In connection with the lysine-arginine interaction reported here it should be mentioned that

Doermann[11] has found the lysine-requiring mutants of Neurospora to be inhibited if the molar concentration of arginine in the medium exceeds that of lysine.

Discussion.—Although it has not been proved directly, there is much evidence which is consistent with the hypothesis that modification of a gene results in a direct change in the activity of one enzyme.[12] On this basis the differences in growth characteristics of the three mutants which behave as alleles would make it appear that different modifications of the same gene are possible, each of which results in a distinct difference in the activity of the enzyme. This view agrees with observations on multiple alleles in other organisms.[13] A study of the properties of the enzyme affected in the three pyrimidineless strains may prove to be useful, but such a study will not be possible until further knowledge of the reaction involved is available.

It would be of interest to know the former condition of the gene which, by mutation, became able to take over the function of the inactive gene at the 3a locus. Perhaps the most obvious possibility is that the wild-type parent carries an inactive duplicate of this locus which was made active by mutation. This possibility would seem to be ruled out, however, by the fact that arginine suppresses the suppressor but does not affect wild type, in which the 3a locus must be active. Also the fact that growth of the mutant strains without the suppressor, on limiting quantities of hydrolyzed nucleic acid, is not affected by arginine would seem to indicate that arginine inhibition of the suppressed mutant is not simply a result of an insufficient supply of pyrimidine.

Two alternative possibilities that have been suggested are, first, that an entirely new gene has arisen from inactive genic material, and, second, that a gene corresponding to an enzyme which promotes a reaction in another synthesis has been modified so that the enzyme now catalyzes the required reaction in pyrimidine synthesis as well. No basis of choice between these two possibilities seems apparent at present.

Summary.—A suppressor, comparable to those found in Drosophila, has occurred in Neurospora. This suppressor acts upon three physiologically different pyrimidine-requiring strains which behave in genetic tests as if they carry allelic genes. It has no effect in combination with two other pyrimidineless strains which are genetically different from these three.

The characteristics of one of the suppressed mutants are sufficiently different from those of wild type to make it unlikely that suppression is due to duplication of the suppressed locus.

* Work supported by grants from the Rockefeller Foundation.

[1] Plough, H. H., *Verh. d. V. Interm. Kongr. f. Vererbung., Berlin I*, **1**, 193 (1928).

[2] Stern, C., *Zeit. f. i. Abst. u. Vererb.*, **52,** 373 (1929).
[3] Bridges, C. B., *Ibid.*, **60,** 207 (1932).
[4] Schultz, J., and Bridges, C. B., *Amer. Nat.*, **66,** 323 (1932).
[5] Beadle, G. W., and Ephrussi, Boris, *Proc. Nat. Acad. Sci.*, **22,** 536 (1936).
[6] Beadle, G. W., *Physiol. Rev.*, **25,** 643 (1945).
[7] Beadle, G. W., and Coonradt, V. L., *Genetics*, **29,** 291 (1944).
[8] Beadle, G. W., and Tatum, E. L., *Amer. Jour. Bot.*, **32,** 678 (1945).
[9] Mitchell, H. K., and Houlahan, M. B., *Fed. Proc.*, in press.
[10] Loring, Hubert S., and Pierce, John G., *Jour. Biol. Chem.*, **153,** 61 (1944).
[11] Doermann, A. H., *Arch. Biochem.*, **5,** 373 (1944).
[12] Beadle, G. W., *Chem. Rev.*, **37,** 15 (1945).
[13] Wright, S., *Physiol. Rev.*, **21,** 487 (1941).

---

*X-RAY AND ULTRAVIOLET STUDIES ON POLLEN TUBE CHROMOSOMES. II. THE QUADRIPARTITE STRUCTURE OF THE PROPHASE CHROMOSOMES OF TRADESCANTIA**

BY C. P. SWANSON

DEPARTMENT OF BIOLOGY, THE JOHNS HOPKINS UNIVERSITY

Communicated May 20, 1947

Most x-ray data dealing with induced chromosome breakage are interpretable on the basis of a divided or an undivided chromosome. The extensive studies on induced breaks in the microspore chromosomes of *Tradescantia*[1] indicate that chromatid and chromosome breaks result, respectively, from changes occurring within a divided or an undivided chromosome. Data from pollen tube chromosomes in the same plant,[2] where the split nature of the chromosomes can be determined at the time of treatment, agree with the above view, and, at the same time, invalidate beyond a doubt the interpretation of Darlington and LaCour[3] as to the origin of induced aberrations of the isochromatid type. The relative lack of mosaics in x-rayed Drosophila[4] suggests that, for the most part, the chromosomes in the sperm head are undivided at the time of treatment, although chromatid breaks are by no means rare in the treated salivary gland chromosomes. It appears, therefore, that, as a general rule, the smallest subdivision of the chromosome to be involved in breakage and rearrangement is the chromatid. This interpretation explains very well similar breakage results obtained from chromosomes treated with ultraviolet,[2] as well as the extensive data of Stadler[5] on endospermal deficiencies in treated maize.

That the chromatid is not the smallest subdivision of the chromosome, however, has long been known (see Nebel[6, 7] and Kuwada[8] for reviews). In addition to earlier observational studies there is much supporting data

of recent date. Mention may be made of the x-rayed chromosomes of *Tradescantia*[2, 9] and *Steatococcus*,[10] the quadripartite nature of the trabant in the microspore chromosomes of *Tradescantia*,[11] and the occurrence of half-chromatid exchanges in the pollen tube chromosomes of the same plant.[2] Carlson[12] has likewise reported the probable occurrence of half-chromatid breaks in the neuroblast chromosomes of *Chortophaga* embryos, as has Marshak[13] in *Vicia*, but these have been dismissed by some[3] as irrelevant observations. Similar half-chromatid breaks are frequent in the pollen tube chromosomes of *Tradescantia* following x-ray or ultraviolet treatment,[2] although they are admittedly difficult to distinguish from the openings between adjacent somatic coils.

Recently a number of aberrations have been found whose interpretation can only be understood in terms of half-chromatid breaks and rearrangements, and it is the purpose of this note to present a consideration of them. Because of their sporadic appearance, no attempt has been made to relate them to dosage, but they are frequent enough to convince the author that the chromosomes in very early prophase (the time of treatment) are quadripartite.

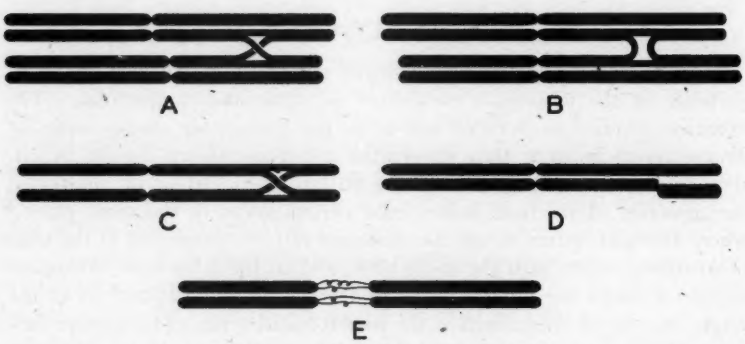

FIGURE 1

Semi-diagrammatic drawings of *Tradescantia* pollen tube chromosomes. A and B. Half-chromatid exchanges between non-homologous chromosomes. C. Half chromatid exchange between the two chromatids of the same chromosome. D. Half-chromatid deletion which has rotated from an inside to an outside position. E. Metaphase chromosome showing a quadripartite condition in the centric region.

Figure 1 illustrates the types of half-chromatid aberrations found. Figures 1A and 1B are exchange breaks. At the points of exchange, the stretched condition of the threads permitted the split in the individual chromatids to be viewed with clarity even though the portions of the chromatids on either side of the break presented a solid and undivided

appearance. Figure 1C illustrates an exchange break between half-chromatids of the same chromosome. The sister chromatids have moved apart from each other at their centromeres, but they are held in close juxtaposition at the point of exchange.

The exchange type of break has never been found following treatment with ultra-violet. The breakage of a single half-chromatid, however, has been observed, and, in figure 1D, the rotation of the broken piece from an inside to an outside position leaves little doubt as to the interpretation: it cannot be confused with the separation of the gyres of a coiled chromatid. Further evidence of the quadripartite structure is illustrated in figure 1E. The chromosomes, in passing down the pollen tube, are frequently stretched, with the strain being evidenced by an attenuation of the centric region. When so stretched, this portion of the chromosome shows four very thin parallel strands, each possessing one or more coils of small diameter. Whether or not these stretched and coiled strands represent the centromere, or a portion of it, is not known; more likely, they are the half-chromatids pulled free from body of the chromatids. If the latter, it may be pointed out that there is no evidence of a chromomeric differentiation even though the threads approximate the thinness of a leptotene chromosome. The coils, too, are of a much smaller diameter than the usual somatic coil.

The fact that the chromatid is further split into half-chromatids in the early prophase stages poses the problem of "What is the functional unit of the chromosome, and what determines its integrity?" In crossing over, genetical and cytological evidence leaves little doubt that the chromatid is that functional unit. The chromatid also appears to be the smallest functional unit in coiling,[13] and, as pointed out above, most x-ray and ultra-violet data are understandable if interpreted in terms of chromatids or chromosomes, but not in terms of some smaller unit. Since, however, the evidence presented reveals a quadripartite structure which under most circumstances behaves as though it were bipartite, the conclusion is unavoidable that there must be some structural unit in the chromosome other than the chromonemata themselves which preserves the integrity of the chromatid, thus enabling it to react as a single entity in crossing over, coiling, breakage, and reattachment. Occasionally this structure can be broken in such a way by radiation as to permit the formation of half-chromatid exchanges, thus revealing a further subdivision of the chromonemata, but under ordinary circumstances the unitary behavior of the chromatid remains. It is tempting to suggest the matrix as the structural unit which determines the behavior of the chromosome, or chromatid, in breakage and reattachment, but the speculative nature of our knowledge concerning the structure and function of the matrix is such as to make one proceed in that direction with caution. It is only possible to state that

the degree to which the chromonemata is subdivided is not the limiting factor which determines the types of breakage and subsequent recombination found following x-radiation. It likewise follows that ultra-violet, although possessing a localized photochemical action, is capable of breaking a bipartite chromatid. Whether it does so by disrupting some structure other than the chromonema itself remains to be determined.

* This research was supported in part by a grant-in-aid from the U. S. Public Health Service.

[1] Sax, K., *Symp. on Quant. Bio.*, IX, 93–103, 1941.
[2] Swanson, C. P., *Jour. Gen. Physiol.*, **26**, 485–494, 1943.
[3] Darlington, C. D., and LaCour, L. F., *J. Gen.*, **46**, 180–267, 1945.
[4] Kaufmann, B. P., *Symp. on Quant. Bio.*, IX, 82–92, 1941.
[5] Stadler, L. J., Proc. 7th Internat. Gen. Congress, 269–276, 1939.
[6] Nebel, B. R., *Bot. Rev.*, **5**, 563–626, 1939.
[7] Nebel, B. R., *Symp. on Quant. Bio.*, IX, 7–12, 1941.
[8] Kuwada, Y., *Cytologia*, **10**, 213–256, 1939.
[9] Nebel, B. R., *Genetics*, **21**, 605–614, 1936.
[10] Hughes-Schrader, Sally, and Ris, H., *J. Exp. Zoo.*, **87**, 429–456, 1941.
[11] Swanson, C. P., unpublished.
[12] Carlson, J. G., *Genetics*, **23**, 596–609, 1938.
[13] Swanson, C. P., *Am. Jour. Bot.*, **30**, 422–428, 1943.

---

*ON THE ASYMPTOTIC SOLUTION OF THE DIFFERENTIAL EQUATION FOR SMALL DISTURBANCES IN A LAMINAR FLOW**

BY WOLFGANG WASOW

SWARTHMORE COLLEGE

Communicated May 29, 1947

The hydrodynamical equation of Orr and Sommerfeld (see C. C. Lin[1] for a detailed analysis) is a special case of differential equations of the form

$$\sum_{j=0}^{4} a_j y^{(4-j)} + \lambda^2 \sum_{k=0}^{2} b_k y^{(2-k)} = 0 \qquad (1)$$

Here a_j, b_k are analytic functions of the complex variable x, λ is a large complex parameter with constant argument, $a_0 = 1$, and b_0 has a zero of first order at some point, say at $x = 0$.

The aim of the present work is to supplement the hydrodynamical theory by a mathematical study of the validity of the formal asymptotic developments used there and, in particular, to shed some additional light on the phenomenon of the so-called "crossing substitutions."

For the convenient statement of our results we need a few definitions. We denote by $Q(x)$ the two-valued function

$$Q(x) = \int_0^x (-b_0(t))^{\frac{1}{2}} dt$$

and by $\eta(x)$ the function

$$\eta(x) = b_0^{-\frac{1}{4}} \exp\left\{\frac{1}{2}\int_0^x \frac{b_0 a_1 - b_1}{b_0} dt\right\}$$

Let S be the domain in the x-plane defined by the inequalities

$$0 < |Q(x)| \leqslant K,$$

where K is a constant chosen so that S does not contain any zeros of b_0. There exist three curves C_1, C_2, C_3, along which $Re(\lambda Q(x)) = 0$. These curves meet at the origin and divide S into three curvilinear sectors S_k, $(k = 1, 2, 3)$. We choose our subscripts so that S_k is bounded by the two curves C_j, $(j \neq k)$, and we assume that C_j belongs to S_k.

We finally introduce a generic symbol $E(S)$ to denote any function of x and λ which, together with its first three derivatives, is bounded uniformly for $|\lambda| \geqslant R > 0$ and for all x in any given closed subdomain of S. The symbols $E(S_k)$, $E(S - C_k)$, etc., are similarly defined.

THEOREM 1. *There exist solutions $A_k(x, \lambda)$, $(k = 1, 2, 3)$ of (1) having the asymptotic representations*

$$A_k(x, \lambda) \doteq e^{\lambda Q(x)}\left[\eta(x) + \frac{E(S - C_k)}{\lambda}\right] \tag{2}$$

Here $Q(x)$ has to be taken with the determination for which $Re(\lambda Q(x)) \leqslant 0$ in S_k, and the functions $Q(x)$ and $\eta(x)$ are regular analytic in $S - C_k$.

Remark: Note that $A_k(x, \lambda)$ is asymptotically equal to two different branches of $e^{\lambda Q(x)}\eta(x)$ on the two sides of C_k.

THEOREM 2. *Let $u(x)$ be any solution of the differential equation $\sum_{k=0}^{2} b_k y^{(2-k)} = 0$. Then there exist solutions $U_k(x, \lambda)$, $(k = 1, 2, 3)$, of (1) such that*

$$U_k(x, \lambda) = u(x) + \frac{E(S - S_k)}{\lambda^2} \tag{3}$$

Both these theorems can be proved by appropriate modifications of methods used for more general differential equations, but for real x only, by W. J. Trjitzinsky.[2]

Theorems 1 and 2 combined give us fundamental systems of solutions of (1) with known asymptotic properties in any two out of the three

sectors S_1, S_2, S_3. The next question is to determine the asymptotic character of the solutions introduced in theorem 2 in the remaining sector S_k. This question is partially answered in the next two theorems:

THEOREM 3. *There exist solutions* $V(x, \lambda)$ *of* (1) *such that*

$$V(x, \lambda) = v(x) + \frac{E(S)}{\lambda^2} \tag{4}$$

where $v(x)$ *is a solution of* $\sum_{k=0}^{2} b_k y^{(2-k)} = 0$ *which is regular at* $x = 0$.

Remark: Note that the essential difference between (3) and (4) is that $E(S)$ is bounded in S and not only in $S - S_k$.

THEOREM 4. *Let* $u(x)$ *be a solution of* $\sum_{k=0}^{2} b_k y^{2-k} = 0$ *which is singular at* $x = 0$, *and let* $U_k(x, \lambda)$ *be a corresponding solution of* (1) *such that* (3) *is true. Then* $U_k(x, \lambda)$ *diverges at every interior point of* S_k, *as* $\lambda \longrightarrow \infty$.

The last two theorems can be derived from theorems 1 and 2 by using three sets of fundamental systems, each having a known asymptotic representation in two of the sectors S_k, and by studying the linear relations between these fundamental systems.

Theorem 4 helps to clarify the questions connected with the so-called "inner friction layers." C. C. Lin and others conclude from the behavior of the solutions which we have called A_1, A_2, that the particular solutions of (1) occurring in the hydrodynamical application approach a discontinuous behavior at C_1 and C_2, as $\lambda \longrightarrow \infty$. Using theorem 4 it is easy to give a complete proof of this statement. Moreover, our results show that C_1 and C_2 are not isolated curves of discontinuity for the limits of these solutions, but that they *diverge in the whole sector* S_3.

It may be remarked, in concluding, that the asymptotic character of $U_k(x, \lambda)$ can be investigated in greater detail, using somewhat different methods. It turns out that $U_k(x, \lambda)$ is in S_k asymptotically equal to a function of the form $\lambda^{-3/2} e^{\lambda Q(x)} \phi(x)$, where $\phi(x)$ is bounded and not identically zero.

* The results presented in this paper were obtained in the course of research conducted under the sponsorship of the Office of Naval Research.

[1] Lin, C. C., "On the Stability of Two-Dimensional Parallel Flows; Parts I–III," *Quarterly Applied Math.*, **3**, 1945.

[2] Trjitzinsky, W. J., "Theory of Linear Differential Equations Containing a Parameter," *Acta Mat.*, **67**, pp. 1–50, 1936.

[3] Wasow, W., "On the Asymptotic Solution of Boundary Value Problems for Ordinary Differential Equations Containing a Parameter," *Jour. Math. Phys.*, **23**, pp. 173–183, 1944.

ABSTRACT GROUP GENERATED BY THE QUATERNION UNITS

BY G. A. MILLER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS

Communicated June 12, 1947

According to Alexander Macfarlane's *Ten British Mathematicians*, page 44 (1916), W. R. Hamilton stated in a letter to his son that on October 16, 1843, he cut with a knife on a stone of Brougham Bridge, near Dublin, Ireland, the fundamental formula with the symbols i, j, k, namely,

$$i^2 = j^2 = k^2 = ijk = -1$$

It is desired to emphasize here the fact that this continued equation is really an equational definition of the abstract non-abelian group of order 8 which is now commonly known as the quaternion group, generated by the quaternion units i, j, k, and that it seems to be the earliest example of an abstract non-abelian group which had been clearly defined then by a set of generational relations between a set of generating operators of the group.

Such definitions soon thereafter became somewhat popular, especially with American writers on group theory, and their origin has frequently been credited to A. Cayley, who emphasized them somewhat later and embodied their fundamental nature in what is known as Cayley's dictum, "a group is defined by the law of combination of its symbols." What A. Cayley embodied in this dictum had really been involved earlier in what W. R. Hamilton regarded as the discovery of quaternions and had noted by the continued equation stated in the preceding paragraph. Probably no other finite non-abelian abstract group has attracted such wide attention and aroused such great hopes as regards its usefulness as did the quaternion group soon after its discovery by W. R. Hamilton in 1843.

A simple proof of the fact that the noted continued equation is really a usable definition of the quaternion group may be given as follows: It results directly from this equation that the two operators i and j are two group operators of order 4 which have a common square and that $ij = k$ and hence that the transform of j by i, that is, $i^3ji = -ki = -j$. In other words, the operator i transforms the operator j into its inverse. The group of order 4 generated by j is therefore transformed into itself by the operator i and involves the square of i. When the group of order 4 generated by j is extended by the operator i there results the well-known group of order 8 which is now commonly called the quaternion group and is characterized by the fact that it contains 6 operators of order 4.

The fact that W. R. Hamilton explicitly laid much stress on the associative law in the combination of operators makes it clearer that he may properly be regarded as the founder of the abstract theory of groups instead

of A. Cayley. On the other hand, W. R. Hamilton did not use the term group but favored the use of his own term quaternions for the development of a theory based upon the special abstract group of order 8 which involves 6 operators of order 4. The associative law was later very commonly embodied in the definitions of an abstract group and the development of these groups did much to emphasize the fundamental importance of this law. These observations in regard to W. R. Hamilton seem to imply that the theory of abstract groups may reasonably be regarded as about ten years older than the assumption that A. Cayley was the founder of this theory would imply. Implicitly, W. R. Hamilton emphasized the importance of the quaternion group in his *Lectures on Quaternions*, page xxvi (1853).

LIMITS FOR THE NUMBER OF SOLUTIONS OF CERTAIN GENERAL TYPES OF EQUATIONS IN A FINITE FIELD

By H. S. Vandiver

Department of Applied Mathematics, University of Texas

Communicated May 22, 1947

Dickson,[1] using cyclotomic integers, obtained an inferior limit for the number of solutions of the congruence

$$x^e + y^e + z^e \equiv 0 \pmod{p} \tag{1}$$

where e and p are given primes with $xyz \not\equiv 0 \pmod{p}$. Hurwitz[2] gave superior and inferior limits for the number of solutions of the congruence

$$ax^e + by^e + cz^e \equiv 0 \pmod{p}, \tag{2}$$

where $abcxyz \not\equiv 0 \pmod{p}$. Recently, the writer has considered generalizations of these results and has arrived at limits both superior and inferior, for the number of solutions of the equation

$$c_1 x_1^{a_1} + c_2 x_2^{a_2} + \ldots + c_s x_s^{a_s} + c_{s+1} = 0 \tag{3}$$

in the x's, where the a's are integers such that $0 < a \leqq p^t - 1$; $s \geqq 2$; the c's belong to a finite field of order p^t, p prime, which will be designated by $F(p^t)$; and

$$c_1 c_2 \ldots c_s x_1 x_2 \ldots x_s \neq 0$$

in $F(p^t)$. As a corollary to this it is possible to show that if we take the c's in (3) as rational integers and put $t = 1$, so that we have a congruence modulo p in effect, then for p sufficiently large the congruence has solutions,

with the x's all prime to p. There is a similar theorem involving algebraic numbers. However, in the present paper we shall confine ourselves to the derivation of certain quadratic relations between the numbers of solutions, x^m, y^m, of trinomial equations of the type

$$1 + ax^m = by^m \tag{3a}$$

for various values of a and b in $F(p^t)$, $abxy \neq 0$, as given in Theorem I. Other quadratic relations connecting such solutions have been given before,[3, 4] but those given here are of a quite different character, and they will be used in another paper in the proofs of the results concerning (3) referred to above.

Consider a finite field of order p^t designated by $F(p^t)$, where p is a prime. Write $p^t = 1 + mc$. Let g be a generator of the cyclic group formed by the non-zero elements of $F(p^t)$. Further let (i, j) denote the number of solutions g^r, g^s of

$$1 + g^{i+rm} = g^{j+sm} \tag{3b}$$

if r and s are each in the range $0, 1, \ldots, c-1$, noting that $g^{mc} = 1$. If we write ind. for index and $g^{\text{ind } x} = x$, represent the index of (-1), modulo m, by ϵ, then for any i and j it is known[4] that

$$(i, j) = (j + \epsilon, i + \epsilon) = (-j + \epsilon, i - j) = (i - j + \epsilon, -j) = (-i, j - i) = (j - i + \epsilon, -i + \epsilon). \tag{4}$$

Also we have[4]

$$\sum_i (i, 0) = c - 1; \quad \sum_i (i, j) = c, \tag{4a}$$

where $i \equiv 0, 1, \ldots, m-1$; $j \equiv 1, 2, \ldots, m-1$ modulo m. We also note that

$$(i, j) = (i + \alpha m, j + \beta m)$$

for any integers α and β. A fundamental relation we shall employ in what follows is (Mitchell,[4] his ψ function defined on p. 165; for $b = a$, $a + b = d$)

$$\psi_{a,d}(\alpha)\psi_{a,d}(\alpha^{-1}) = p^t \tag{4b}$$

where

$$\psi_{a,d}(\alpha) = \sum_{i,j} (i, j)\alpha^{-ai+dj}$$

α being a primitive mth root of unity, a and d any integers subject to the conditions $a \not\equiv 0$, $d \not\equiv 0$, $a - d \not\equiv 0 \pmod m$ with i and j ranging independently over the integers $0, 1, \ldots, m-1$.

We shall now determine the quantities

$$\sum_{i,j}^{0 \text{ to } m-1} (i, j)\,(i + h, j + k) = A_{hk} \tag{5}$$

First we determine A_{00}. Now if $a \not\equiv 0 \pmod m$ $d \not\equiv 0 \pmod m$, $d - a \not\equiv 0 \pmod m$ then if α is a primitive mth root of unity, (4b) gives

$$\sum_{hijk}^{0 \text{ to } m-1} (i, j)\,(i + h, j + k)\alpha^{-ah+dk} = p^t \tag{6}$$

Let $p^t = 1 + mc$. Consider the summation

$$\sum_{d=0}^{m-1} \sum_{hijk}^{0 \text{ to } m-1} (i, j)(i + h, j + k)\alpha^{-ah+dk} = \sum_{d=0}^{m-1} C_{ad} \tag{7}$$

For $a \not\equiv 0 (\text{mod } m)$, $d \not\equiv 0 (\text{mod } m)$, $a - d \not\equiv 0 (\text{mod } m)$ each term in the sum with respect to $d = p^t$. Now consider the value when $d = 0$, which is

$$\sum_{hijk}^{0 \text{ to } m-1} (i, j)(i + h, j + k)\alpha^{-ah}$$

Set $i + h = i'$, $j + k = j'$ then this expression becomes

$$\sum_{i,j,i',j'} (i, j)(i', j')\alpha^{-(i'-i)a} = C_{ao} \tag{8}$$

This is obviously equal to

$$\left(\sum_{i,j}^{0 \text{ to } m-1} (ij)\alpha^{ia}\right)\left(\sum_{i',j'}^{0 \text{ to } m-1} (i'j')\alpha^{-i'a}\right) \tag{8a}$$

and each of these equals (-1) if $a \not\equiv 0 (\text{mod } m)$, hence

$$C_{ao} = 1 \tag{9}$$

Now consider the case when $a - d \equiv 0 (\text{mod } m)$ in (7). The corresponding term reduces to

$$\sum (i, j)(i', j')\alpha^{-a(h-k)} = C_{aa},$$

or

$$\sum (i, j)\alpha^{-a(i-j)} \sum (i', j')\alpha^{-a(i'-j')}$$

If $i - j = f$, then for $a \not\equiv 0 (\text{mod } m)$, using (4),

$$\sum_{j,f} (f + \epsilon, -j)\alpha^{-af} = \sum_{j,f} (f, -j - \epsilon)\alpha^{-af} = -1$$

and similarly for the second factor. Hence

$$C_{aa} = 1, a \not\equiv 0 (\text{mod } m). \tag{10}$$

Now if $a \equiv 0 \pmod{m}$ in (8) then we find

$$C_{00} = (mc - 1)^2 \tag{11}$$

employing (8a) and (4a).

We now simplify (7) under the assumption that $a \not\equiv 0 \pmod{m}$. Using (7), (8), (9) and (10), we find

$$\sum_{d=0}^{m-1} C_{ad} = (m - 2)p^t + 2; \quad a \not\equiv 0 \pmod{m}. \tag{12}$$

For the case where $a \equiv 0 \pmod{m}$ we obtain in the same way that (9) was derived

$$C_{od} = 1, \quad d \not\equiv 0 \pmod{m}. \tag{13}$$

Hence

$$\sum_{d=0}^{m-1} C_{od} = (mc - 1)^2 + m - 1 \tag{13a}$$

Now the left hand member of (12) may be written

$$\sum_{d=0}^{m-1} \sum_{hk} A_{hk}(\alpha)^{-ah+dk}$$

where A_{hk} is defined as in (5).
Now

$$\sum_{d=0}^{m-1} \sum_{hk} A_{hk}\alpha^{-ah+dk}$$

$$= \sum_{hk} A_{hk}\alpha^{-ah}(1 + \alpha^k + \ldots + \alpha^{(m-1)k})$$

$$= 0 \text{ if } k \not\equiv 0 \pmod{m}.$$

Hence

$$\sum_{d=0}^{m-1} \sum_{hk} A_{hk}\alpha^{-ah+dk} = mA_{ho}\alpha^{-ah},$$

or from (12),

$$m\sum_{h} A_{ho}\alpha^{-ah} = (m - 2)p^t + 2, \tag{14}$$

if $a \not\equiv 0 \pmod{m}$ and it equals $(mc - 1)^2 + m - 1$ otherwise.
Hence

$$\sum_{a=0}^{m-1} mA_{ho}\alpha^{-ah} = (m - 2)(mc + 1)(m - 1) + 2(m - 1) + (mc - 1)^2 + m - 1$$

or

$$m^2A_{oo} = m^2c^2 + m^3c - 3m^2c + m^2$$

or

$$A_{oo} = (c - 1)^2 + (m - 1)c \tag{14a}$$

Now for $a \not\equiv 0 \pmod m$ we have from (14)

$$m\sum_h A_{ho}\alpha^{-ah+sa} = \alpha^{sa}((m - 2)p^t + 2)$$

and for $a \equiv 0 \pmod m$,

$$m\sum_h A_{ho} = (mc - 1)^2 + m - 1.$$

Hence

$$m\sum_{a=0}^{m-1}\sum_h A_{ho}\alpha^{-ah+sa} = (mc - 1)^2 + m - 1 - ((m - 2)p^t + 2) = (c^2 - c)m^2,$$

or

$$m^2A_{ho} = m^2(c^2 - c),$$

$$A_{ho} = c^2 - c,\ h \not\equiv 0 \pmod m. \tag{15}$$

We shall now show also that

$$A_{ok} = c^2 - c,\ k \not\equiv 0 \pmod m. \tag{16}$$

We have

$$A_{ok} = \sum_{i,j}(i, j)(i, j + k)$$

But by (1)

$$(i, j) = (j + \epsilon,\ i + \epsilon)$$

$$(i, j + k) = (j + k + \epsilon,\ i + \epsilon)$$

Hence

$$\sum_{i,j}(i, j)\,(i, j + k) = \sum_{j+\epsilon,\ i+\epsilon}((j + \epsilon) + k,\ i + \epsilon)\,(j + \epsilon,\ i + \epsilon),$$

where $j + \epsilon$ and $i + \epsilon$ range over the same set $0, 1, \ldots m - 1$, modulo m as do i and j so that this equation gives (16), using (15).

We lastly determine A_{hk} with $h \not\equiv 0$, $k \not\equiv 0 \pmod m$. Consider the sum

$$\sum_{d=0}^{m-1}\ \sum_{hijk\ \bullet}^{0\text{ to }m-1}(ij)(i + h\,j, + k)\alpha^{-ah+dk-vd} = B_{adv} \tag{17}$$

We know from (6) that each term corresponding to a given d equals $p^t\alpha^{-vd}$ except when $d \equiv 0$, $a \equiv 0$, or $d \equiv a_1 \pmod m$. For $d \equiv 0$, then

the corresponding term equals unity if $a \not\equiv 0 \pmod{m}$, using (9). For $d \equiv a$ we have by (10) that the term is α^{-av}. Hence

$$B_{adv} = 1 + \alpha^{-v}p^t + \alpha^{-2v}p^t + \ldots + \alpha^{-(m-1)v}p^t + \alpha^{-av} - \alpha^{-av}p^t$$

$$B_{adv} = 1 - p^t + \alpha^{-av}(1 - p^t) \tag{18}$$

But B_{adv} can also be written as

$$mA_{hv}\alpha^{-ah} + \sum_{\substack{k=0 \\ k\neq v}}^{m-1} A_{hk}\alpha^{-ah}(1 + \alpha^{k-v} + \ldots + \alpha^{(m-1)(k-v)})$$

and the last summation is 0. Hence by (18)

$$mA_{hv}\alpha^{-ah} = 1 - p^t + \alpha^{-av}(1 - p^t)$$

or

$$A_{hv}\alpha^{-ah} = -c - \alpha^{-av}c \tag{19}$$

and

$$A_{hv}\alpha^{-ah+sa} = -c\alpha^{sa} - c\alpha^{-av+sa} \tag{20}$$

We also have, from (17), for $a \equiv 0 \pmod{m}$

$$B_{odv} = mc^2 - 2c \tag{21}$$

for if we take (17) with $a \equiv 0 \pmod{m}$ then for $d \equiv 0 \pmod{m}$, we find that the corresponding term is $(mc - 1)^2$ while the other terms, corresponding to each d are $\alpha^{-v}, \alpha^{-2v}, \ldots, \alpha^{-(m-1)v}$, using (13). From (20) and (21) we find

$$\sum_{a=0}^{m-1} A_{hv}\alpha^{-ah+sa} = \sum_{a=0}^{m-1}(-c\alpha^{av} - c\alpha^{-av+sa}) + mc^2 - 2c = mc^2,$$

if $v \not\equiv s \pmod{m}$, or

$$A_{sv} = c^2; \quad s \not\equiv 0, \; v \not\equiv 0, \; s \not\equiv v \pmod{m}. \tag{21a}$$

For $v = s$ we have from the above

$$A_{ss} = c^2 - c; \quad s \not\equiv 0 \pmod{m}. \tag{22}$$

Hence we have, employing (14a), (15), (16), (21a) and (22),

THEOREM I. *Set*

$$A_{hk} = \sum_{i,j}^{0 \text{ to } m-1} (i, j)(i + h, j + k),$$

where (i, j) is the number of sets of values f and l in the set $0, 1, \ldots, c - 1$, which satisfy the equation

$$1 + g^{i+fm} = g^{+m}$$

in the finite field $F(p^t)$ *where* p *is a prime such that* $p^t = 1 + cm$, g *being a primitive root in* $F(p^t)$. *Then*

$$A_{oo} = (c - 1)^2 + c(m - 1); \quad (23)$$

$$A_{ho} = A_{ok} = c^2 - c, \; h \not\equiv 0, \; k \not\equiv 0 \pmod{m}; \quad (24)$$

$$A_{hk} = c^2, \text{ with } h \not\equiv k \pmod{m}; \; h \not\equiv 0, \; k \not\equiv 0 \pmod{m}; \quad (25)$$

$$A_{hk} = c^2 - c, \quad (26)$$

if $h \equiv k \pmod{m}$.

[1] *Crelle*, **135**, 181–188 (1909). Cf. also Pellet, *Bull. Math. Soc. France*, **15**, 80–93 (1886).

[2] *Crelle*, **136**, 272–292 (1909).

[3] THESE PROCEEDINGS, **32**, 47–52 (1946).

[4] Mitchell, *Proc. Amer. Math. Soc.*, **17**, 167 (1916).

CHARACTERISTIC COÖRDINATES FOR HYPERBOLIC DIFFERENTIAL EQUATIONS IN THE LARGE

BY T. Y. THOMAS

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY

Communicated May 31, 1947

Partial differential equations of hyperbolic type are habitually treated by the introduction of *characteristic coördinates*. When such equations occur in physical problems their solutions in the large are necessarily demanded. In spite of this fact the existence of characteristic coördinates appears to be established only locally. We shall here give conditions for the existence of a (1, 1) differentiable transformation defining the characteristic coördinate system *in the large*. The form of these conditions is suitable for practical application but such applications of the theorem will not be included in this communication.

Let D be an open simply connected two dimensional domain, finite or infinite, referred to a system of rectangular coördinates x^α. At each point of D characteristic directions λ are defined as the solutions of the equation $\Sigma g_{\alpha\beta}\lambda^\alpha\lambda^\beta = 0$ in which the summation is over the values 1, 2 of the indices and it is supposed that the coefficients $g_{\alpha\beta}$ are continuous and have continuous first partial derivatives in D. Assuming det. $|g_{\alpha\beta}| < 0$ (hyperbolic case) there will be exactly two distinct characteristic directions at each point of D and these directions will generate two families or congruences of characteristic curves each of which will cover D completely. In fact any point P of D is contained in a local coördinate system, e.g., a

system obtained by rotation of the x coördinate neighborhood of P, within which $g_{22} \neq 0$ and relative to this system the characteristic curves of either congruence are given as solutions of a differential equation of the form $dy/dx = f(x, y)$ where the function $f(x, y)$ is continuous and has continuous partial derivatives with respect to x and y. Each characteristic curve therefore has a parametric representation $x^\alpha(s)$, admitting continuous first derivatives with respect to the parameter s which can be taken as the arc length, and such that $(dx^1/ds)^2 + (dx^2/ds)^2 \neq 0$ at any point of the curve. For brevity such a curve is said to be *regular*. Evidently a characteristic curve cannot end at a point of D (interior point); hence a characteristic curve can be continued indefinitely in either direction through a point of D or it can be continued until it meets a boundary point of D at which point the curve will be considered to terminate.

Now we consider a regular curve C in D defined parametrically by $x^\alpha(p)$ for $a < p < b$ where either a or b can be finite or infinite. The curve C will be said to be a *cross section* for a congruence of curves covering D under the following conditions: (1) each curve of the congruence meets C in one and only one point and (2) a curve of the congruence is not tangent to C at its point of intersection with C. We now make the following fundamental assumption. *Each of the above congruences of characteristic curves in D admits a cross section.*

Denoting one of the characteristic congruences by A and the other by B, let C_A and C_B be the cross sections of A and B. Let C_A and C_B have the parametric representations $x^\alpha(p)$ with $a < p < b$ and $x^\alpha(q)$ with $c < q < d$, respectively. Let $h(p)$, $a \leqq p \leqq b$, be a monotorically increasing (or decreasing) and continuously differentiable function of p and $k(q)$, $c \leqq q \leqq d$, an analogous function of the variable q. Now define the function $\phi(x)$ over D by the condition that $\phi(x)$ be constant along the individual curves of the congruence A and that it assumes a value on any particular curve of A equal to the value of the function $h(p)$ for the value of the parameter p at the point where the curve cuts C_A. In a corresponding way we define a function $\psi(x)$ over D using the characteristic congruence B, the cross section C_B and the function $k(q)$. It is evident that both functions $\phi(x)$ and $\psi(x)$ are continuous and bounded over D.

We now prove that the functions $\phi(x)$ and $\psi(x)$ have continuous first partial derivatives in D. Neighboring any point Q of D we can make a continuous differentiable transformation, $x \leftrightarrow y$, to a y coördinate system relative to which the A and B curves are the parametric lines y^1 and y^2, respectively (local characteristic coördinate theorem). Hence in the y system $\phi = \phi(y^2)$ and $\psi = \psi(y^1)$. Two such systems y and $\bar{y}$ are evidently related by a transformation of the form $y^1 = f(\bar{y}^1)$, $y^2 = g(\bar{y}^2)$ in their common domain where the functions f and g have continuous non-vanishing derivatives. Now let Q be any point in D and consider the A curve

through Q which intersects C_A. Denote this point of intersection by P. The arc PQ of the A curve can be covered by a finite number m of the above local coördinate systems. Denote these systems by $y_1, \ldots, y_m$ and suppose that system y_1 contains the point P, system y_m the point Q, and furthermore that any two successive systems y_i, y_{i+1} contain a region of intersection so that it is possible to traverse the arc PQ by passing through these local systems in the order $y_1, \ldots, y_m$. Since $\phi = \phi(y_m{}^2)$ in the y_m system the derivative $\partial\phi/\partial y_m{}^1$ exists at Q and is equal to zero. It suffices therefore to prove the existence and continuity of the derivative $d\phi/dy_m{}^2$ at Q. For this purpose we write

$$\frac{\Delta\phi}{\Delta y_m{}^2} = \frac{\Delta h}{\Delta p}\frac{\Delta p}{\Delta y_1{}^2}\frac{\Delta y_1{}^2}{\Delta y_2{}^2}\cdots\frac{\Delta y_{m-1}{}^2}{\Delta y_m{}^2}. \tag{1}$$

The last factor in the right member of (1) can be considered as the ratio of the increments $\Delta y_{m-1}{}^2$ and $\Delta y_m{}^2$ at a point on the arc PQ in the intersection of the y_{m-1} and y_m coördinate regions. A similar remark applies to preceding ratios up to and including the ratio $\Delta y_1{}^2/\Delta y_2{}^2$ where the increments are taken at a point on the arc PQ in the intersection of the y_1 and y_2 regions. These ratios approach limits since the local coördinate transformations involved are differentiable. In the y_1 system the curve C_A is represented by $y_1{}^\alpha(p)$. Hence $dy_1{}^2/dp$ exists and this derivative is different from zero since the curves $y_1{}^2 = \text{const.}$ are not tangent to C_A. Hence the inverse relation $p = p(y_1{}^2)$ exists and is differentiable. This proves the existence of the limit of $\Delta p/\Delta y_1{}^2$. Finally the limit $\Delta h/\Delta p$ exists since the function $h(p)$ is differentiable by hypothesis. We thus deduce the existence of the derivative $d\phi/dy_m{}^2$ by taking the limit of the right member of (1) as $\Delta y_m{}^2 \to 0$; moreover, this derivative is seen to be continuous since it is expressible as the product of derivatives each of which is continuous. The existence and continuity of the derivatives of the function $\phi(x)$ with respect to the original variables x^α follows from the fact that the local transformation $x \leftrightarrow y_m$ possesses continuous partial derivatives. In a similar manner we prove the existence and continuity of the partial derivatives of the function $\psi(x)$.

The derivative $d\phi/dy_m{}^2$ is not equal to zero at Q since each of the derivatives which results by taking the limit of the right member of (1) is observed to be different from zero. Similarly $d\psi/dy_m{}^1$ is different from zero at Q. Hence, the functional determinant of ϕ, ψ relative to the local y system is not equal to zero. From this it follows that the functional determinant relative to the original coördinates x^α is not equal to zero at any point Q of D since this determinant is a scalar density under coördinate transformations (the functions ϕ and ψ being absolute scalars). The transformation $\phi = \phi(x)$, $\psi = \psi(x)$ is therefore (1, 1) locally.

But the above transformation $\phi(x), \psi(x)$ is easily seen to be (1, 1) *in the large, i.e., over the entire domain D.* For suppose two distinct points P and Q of D are transformed into the same point (ϕ, ψ). Then P and Q must lie on the same A curve and also on the same B curve. Introducing the above mentioned parameter s for the A curve we see that for some point M on this curve between P and Q we must have

$$\Sigma \frac{\partial \phi}{\partial x^\alpha} \frac{dx^\alpha}{ds} = 0, \ \Sigma \frac{\partial \psi}{\partial x^\alpha} \frac{dx^\alpha}{ds} = 0;$$

the first of these equations follows since ϕ = const. along the A curve and the second from the mean value theorem and the assumption that the function ψ has the same value at points P and Q. Since not both derivatives dx^1/ds and dx^2/ds vanish simultaneously the determinant of the coefficients $\partial\phi/\partial x^\alpha$ and $\partial\psi/\partial x^\alpha$ in the above equations must vanish at the point M. But this contradicts the fact, already proved, that this determinant is nowhere zero in D. This proves the above italicized statement.

Now $\Sigma(\partial\phi/\partial x^\alpha)\lambda^\alpha = 0$ along any A curve. Hence the vector having components $\partial\phi/\partial x^\alpha$ is perpendicular to the vector λ which generates the A curves. Putting $\epsilon_{\alpha\beta} = \sqrt{-g}\, e_{\alpha\beta}$, where $e_{11} = e_{22} = 0$, $e_{12} = -e_{21} = 1$ and g = det. $|g_{\alpha\beta}|$, we can therefore write $\partial\phi/\partial x^\alpha = \gamma\epsilon_{\alpha\mu}\lambda^\mu$ where γ is a proportionality factor. Hence

$$\Sigma g^{\alpha\beta} \frac{\partial \phi}{\partial x^\alpha} \frac{\partial \phi}{\partial x^\beta} = \gamma^2 \Sigma g^{\alpha\beta} \epsilon_{\alpha\mu} \epsilon_{\beta\nu} \lambda^\mu \lambda^\nu = \gamma^2 \Sigma g_{\mu\nu} \lambda^\mu \lambda^\nu = 0.$$

There is a similar equation for the function $\psi(x)$. It follows that $g^{11} = g^{22} = 0$ relative to the (ϕ, ψ) coördinate system. This condition can also be expressed by writing $g_{11} = g_{22} = 0$. The components g^{12} and g_{12} must be everywhere different from zero relative to the (ϕ, ψ) system since det. $|g_{\alpha\beta}| < 0$ over D.

The results obtained can be summarized by the following theorem. *Let D be an open simply connected domain, finite or infinite, which is referred to a system of rectangular coördinates x^α. Suppose that each of the two characteristic congruences defined over D by the equation $\Sigma g_{\alpha\beta}\lambda^\alpha\lambda^\beta = 0$ admits a cross section, where det. $|g_{\alpha\beta}| < 0$ and the coefficients $g_{\alpha\beta}$ are continuous and have continuous first partial derivatives in D. Then there exist functions $\phi(x)$ and $\psi(x)$ which are continuous and have continuous first partial derivatives over D such that (i) these functions define a (1,1) map of D into a bounded region of the two dimensional number space of coördinates ϕ, ψ and (ii) the components $g_{11} = g_{22} = 0$ when taken relative to the (ϕ, ψ) system.* The coördinates ϕ, ψ thus defined are called *characteristic coördinates.* Actually our demonstration of this result has not made explicit use of the condition that D be simply connected. However it is evident that the region covered by the ϕ, ψ coördinate system is simply connected; from

this fact and the continuity of the map we infer the simple connectivity of the domain D.

Remark 1. In certain cases the conditions of the above theorem will not apply to a particular domain D but will apply if the domain is suitably extended. For example consider the infinite strip D between the lines $x = \pm 1$. If one of the congruences consists of the segments of parabolas $y = (x - a)^2$ contained in D, where a is an arbitrary parameter, the y axis can be taken as the cross section. However if D is limited to the interior of the square $x = \pm 1$, $y = \pm 1$ or even to the infinite region between the lines $x = \pm 1$ and above the line $y = 1$, it is evident that no cross section is possible without the identification of certain segments of parabolas. In practical applications, however, it is expected that one will be in a position to choose the appropriate domain D for which the simplified conditions of the theorem, as stated, will apply.

Remark 2. If we do not impose the condition that the congruences of characteristic curves admit cross sections it can be proved that in any simply connected region R whose closure $\bar{R}$ is contained in D functions $\phi(x)$ and $\psi(x)$ can be defined which (1) are continuous and have continuous first partial derivatives in R and (2) define a locally (1, 1) transformation to characteristic coördinates. But now the transformation is not necessarily (1, 1) over the entire region R.

PHYSICAL CURVES

By Edward Kasner

Department of Mathematics, Columbia University, New York

Communicated May 27, 1947

1. We consider the motion of a particle in the plane under the action of any positional field of force. The general equations of motion are

$$\frac{d^2x}{dt^2} = \phi(x, y), \quad \frac{d^2y}{dt^2} = \psi(x, y), \tag{1}$$

where (ϕ, ψ) are the rectangular components of the force acting at any point (x, y), and the mass of the particle is assumed to be unity.

The differential equation of third order representing the system of trajectories, found by eliminating the time t from (1), is

$$(\psi - y'\phi)y''' = [\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2]y'' - 3\phi y''^2. \tag{2}$$

This is not an arbitrary differential equation of third order. Hence the

system of ∞^3 trajectories generated by a field of force must have peculiar geometric properties.

I have developed the differential geometric aspects of dynamics in the Princeton Colloquium.[1] A completely characteristic set of five properties of the system of ∞^3 trajectories defined by the differential equation (2) was discussed and I obtained also many other equivalent characteristic sets.[2]

Of importance in dynamics are the ∞^1 lines of force. These are defined by the differential equation of first order

$$dy/dx = y' = \psi/\phi \tag{3}$$

This family has no peculiar geometric properties since any family of ∞^1 curves can represent the lines of force of some positional field.

2. In connection with a field of force, the only curves usually studied are the lines of force and the trajectories. However, other noteworthy systems of curves are connected with the field, for example, brachistochrones, catenaries, velocity curves.

All these systems, together with the trajectories, may be obtained as special cases of one simple general physical problem: to find curves along which a constrained motion is possible such that the pressure is proportional to the normal component of the force.

3. If an arbitrary curve is drawn in the plane field of force, and the particle (of unit mass) is started along it from one of its points with a given speed, the constrained motion along the given curve is determined. The acceleration along the curve is given by T, the tangential component of the force vector. So the speed at any point is determined by

$$v^2 = 2\int T ds. \tag{4}$$

The pressure P (of course, normal to the curve since the curve is considered smooth) is given by the elementary formula

$$P = \frac{v^2}{r} - N, \tag{5}$$

where N is the normal component of the force vector and r is the radius of curvature of the curve. If we increase the initial speed, the effect is to increase v^2 by a constant c; and hence P changes by the addition of a term of the form c/r.

4. If the given curve is a trajectory, the initial speed may be so chosen that the pressure vanishes throughout the motion; that is, trajectories may be defined as curves of no constraint. Of course, if a different initial speed is used, P will be of the form c/r; but as regards the curves, they are completely characterized by $P = 0$.

If the given curve is a brachistochrone and if the motion along it is brachistochronous (minimum time), Euler proved (assuming that the force

is conservative) that the pressure is double the normal component of the acting force and opposite to it in direction, that is, $P = -2N$. If the force is not conservative, the real brachistochrones, as defined by a problem of the calculus of variations, forms a quadruply infinite system. The curves defined by the property $P = -2N$ then form a triply infinite system of what should be called pseudo-brachistochrones. These curves are actually brachistochrones only in the conservative case. No ambiguity, however, will arise by terming the system here considered brachistochrones instead of pseudo-brachistochrones.

5. *The general physical problem suggested is to find curves such that P shall be proportional to N, that is,* $P = kN$. To a given value of k, there correspond ∞^3 such curves: the system so obtained will be denoted by S_k.

The four cases of main physical interest are as follows:

$k = 0$ gives S_0, the system of *trajectories;*

$k = -2$ gives S_{-2}, the system of *brachistochrones;*

$k = 1$ gives S_1, the system of *catenaries;*

$k = \infty$ gives S_∞, the system of *velocity curves.*

The last case requires a justification in terms of limits which is easily carried out analytically.

The third case follows from the known fact that when an inextensible flexible homogeneous string is suspended in any field of force, the resulting form of equilibrium, called a catenary in the general sense of the term, has the dynamical property that, when a particle, started out with the proper initial velocity, rolls along the curve, the pressure at any point equals the normal component of the force; that is, catenaries are defined by $P = N$ corresponding to $k = 1$.

6. Of course a triply infinite system S_k exists for any value of the parameter k. The differential equation of the system, in intrinsic form, is obtained by eliminating v from the equations

$$v^2/r = (k + 1)N, \quad vv_s = T. \tag{6}$$

The result is

$$Nr_s = nT - rN_s, \tag{7}$$

where $n = 2/(k + 1)$.

The differential equation of third order defining the system S_k of ∞^3 curves is explicitly

$$(\psi - y'\phi)y''' = [\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2]y'' - \left[3\phi + \frac{(n-2)(\phi + y'\psi)}{1 + y'^2}\right]y''^2. \tag{8}$$

This obviously reduces to the familiar trajectory equation (2) when $n = 2$

corresponding to $k = 0$. Brachistochrones correspond to $n = -2$, catenaries to $n = 1$, velocity curves to $n = 0$. From (8) we derive the fundamental theorem of section 9.

7. Now we state the characteristic properties of a system S_k of the above type for any value of n, that is, any value of k.

Property 1. For any given lineal element (x, y, y'), the foci of the osculating parabolas of the single infinity of curves determined by the given element lie on a circle passing through the given point.[3]

Property 2. At any point 0, the tangent of the angle which the focal circle makes with the given element is to the tangent of the angle which the given element makes with a certain fixed direction at 0 (the direction of the acting force) as 3 is to $n + 1$, that is, as $3k + 3$ is to $k + 3$.

Property 3. Through a given point, there pass a single infinity of curves admitting hyperosculating circles of curvature; the centers of these circles lie on a conic passing through the given point in the direction of the force vector.

Property 4. The normal at the given point 0 cuts the conic described in Property 3, at a distance equal to $n + 1$, that is $(k + 3)/(k + 1)$, times the radius of curvature of the line of force passing through 0.

Property 5. This is of the same form as Property V obtained in the discussion of trajectories, the number 3 being replaced by the number $n + 1$. When the point 0 is moved, the associated conic referred to above changes in the following manner. Take any two fixed perpendicular directions for the x direction and the y direction; through 0 draw lines in these directions meeting the conic again at A and B, respectively. Also construct the normal at 0 meeting the conic again at N. At A draw a line in the y direction meeting this normal in some point A' and at B draw a line in the x direction meeting the normal in some point B'. The variation property referred to takes the form (where $\omega = \psi/\phi$ = slope of force vector).

$$\frac{\partial}{\partial x}\frac{1}{AA'} + \frac{\partial}{\partial y}\frac{1}{BB'} + \frac{\omega\omega_{xy} - \omega_x\omega_y}{(n+1)\omega^2} = 0. \qquad (9)$$

[3] See the diagram on page 11 of the Princeton Colloquium.

8. While the properties corresponding to different values of k are analogous, they are, of course, not identical. The first property is common to all the systems. But the second property involves the parameter k. Thus while for trajectories the constant ratio that appears is 1 (bisection), it is -3 for brachistochrones, 3/2 for catenaries, and 3 for velocity curves. Not only are the triply infinite systems S_k, corresponding to different values of k, distinct in any given field of force, but also no two systems arising in two distinct fields can ever coincide. For example, if a certain system of ∞^3 curves arises as trajectories in one field, it cannot also arise as catenaries in either the same or another field.

9. FUNDAMENTAL THEOREM. *It is found that if a curve of the system S_k starts in the direction of the line of force, the curvature of this curve is, in general, zero. But for one special curve, it is found that the ratio ρ of the curvature of the curve to the curvature of the line of force is*[4]

$$\rho = \frac{1}{n+1} = \frac{k+1}{k+3}. \tag{10}$$

For trajectories the special curve is obtained by starting the particle from *rest*, and the ratio is $\rho = {}^1/_3$ as in our original theory. For brachistochrones $\rho = -1$, for catenaries $\rho = {}^1/_2$, for velocity curves $\rho = 1$.

10. If we combine all the systems S_k in a given field of force, we obtain a quadruply infinite system which we now proceed to study. The differential equation of the fourth order defining this system is obtained by eliminating k from the equation of S_k. We shall not write this differential equation of fourth order but state some of the geometric properties.

For any given curvature element (x, y, y', y''), the ∞^1 curves of the system have the property that the locus of the third center of curvature is a parabola with axis parallel to the fixed radius of curvature, that is, perpendicular to the initial direction y'. However, the best statement is: *If for each of the curves, we construct the osculating conic (five-point contact), the locus of the centers of these conics is a conic passing through a given point in the given direction.*[5]

11. Appell showed that any collineation carries the ∞^3 dynamical trajectories of a field of force into the trajectories of a new field of force. We have shown that the collineation group is the largest group of transformations with this property, not only for point transformations, but also for contact transformations.

Goursat proved that a conformality converts the dynamical trajectories of a conservative field of force into the trajectories of a conservative field. Our final result is that the conformal group of transformations consists of all the correspondences that transforms the trajectories of a conservative field of force into the trajectories of a conservative field. This applies not only to the ∞^2 curves for a given energy constant (natural family), but also to the total system of ∞^3 trajectories.

The extension of the transformation theories of Appell and Goursat to catenaries and all systems S_k will be studied elsewhere.

[1] Kasner, Edward, "Differential-Geometric Aspects of Dynamics," Amer. Math. Soc. Colloquium Publications, vol. 3, 1913, 1934 (referred to as Princeton Colloquium). Also see a series of papers in *Trans. Amer. Math. Soc.*, vols. 7–11 (1906–1910). Tautochrones (isochronous motion) are not included among systems S_K of the present paper.

[2] Recently Terracini has given an alternate projective characterization of the ∞^3 dynamical trajectories. See "sobre la ecuacion diferencial $y''' = G(x, y, y')y'' + H(x, y, y')y''^2$," *Revista de Matematicas*, 2, 245–329 (1941), Tucuman, Argentina.

[3] Kasner, Edward, and De Cicco, John, "A Generalized Theory of Dynamical Trajectories," *Trans. Amer. Math. Soc.*, **54**, 23–38 (1943).

[4] Kasner, Edward, and Mittleman, D., "A General Theorem on the Initial Curvature of Dynamical Trajectories," *Proc. Nat. Acad. Sci.*, **28**, 48–52 (1942). Also "Extended Theorems in Dynamics," *Science*, **95**, 249–250 (1942). Also a long paper appearing in the current volume of *Revista de Matematics*.

[5] Kasner, Edward, "Systems of Extremals in the Calculus of Variations," *Bull. Amer. Math. Soc.*, **13**, 290 (1907).